



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶ : G06F 17/27	A1	(11) International Publication Number: WO 99/14689 (43) International Publication Date: 25 March 1999 (25.03.99)
(21) International Application Number: PCT/US98/19432 (22) International Filing Date: 17 September 1998 (17.09.98) (30) Priority Data: 08/932,937 17 September 1997 (17.09.97) US (71) Applicant: UNISYS CORPORATION [US/US]; Township Line and Union Meeting Roads, P.O. Box 500, Blue Bell, PA 19424-0001 (US). (72) Inventors: KENDALL, Daythal, Lee; 416 Roslyn Avenue, Glenside, PA 19038 (US). WADSWORTH, Dennis, Lee; 101 Hoffman Circle, Downingtown, PA 19335 (US). BOUZID, Ahmed, Twefik; 70 Overlook Drive, Norristown, PA 19401 (US). DAHL, Deborah, Anna; 1820 Gravers Road, Norristown, PA 19401 (US). HUA, Hua; 501 Cindy Circle, Penllyn, PA 19422 (US). (74) Agent: STARR, Mark; Unisys Corporation, Township Line and Union Meeting Roads, P.O. Box 500, Blue Bell, PA 19424-0001 (US).		(81) Designated States: CA, CN, IL, JP, SG, European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE). Published <i>With international search report.</i> <i>Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i>

(54) Title: SYSTEM AND METHOD FOR CREATING A LANGUAGE GRAMMAR

(57) Abstract

The present invention is a computer software system that allows the developer of a speech-enabled system to create a grammar and corpus for use in the system. A table interface is used, and phrases in the grammar are entered into cells in the table. The table also includes token data which corresponds to each valid utterance. When the grammar is defined, the computer software system automatically traverses the table to enumerate all possible valid utterances in the grammar. This traversal generates a listing (corpus) of valid utterances and their respective tokens. This listing can then be used to interpret spoken utterances for a speech-enabled system. The computer software system also transcribes the grammar rules found in the table to a format compatible with a variety of supported commercially-available speech recognizers.

The screenshot shows a software window titled "TestProject-Speech Assistant Tool". It has a menu bar with "File", "Compartment", and "About". Below the menu is a "Compartment List" dropdown. The main area is titled "Compartment: COMPARTMENT1" and contains a table editor with tabs for "File", "Table", and "Window". The "Table" tab is active, showing a table with 12 rows and 7 columns. The table is titled "BALANCE" and contains the following data:

	1	2	3	4
1	my	checking	balance	
2		savings	balance	
3	the balance in my	checking	account	
4		savings	account	
5		checking	account	

Below the table editor is a section titled "Starting Table: MAIN" which displays a list of generated utterances and their corresponding tokens:

	1	2	3
1	I think	I need	LOAN
2		I want	
3		I'd like	a
4		I would like	the term to be
5			it for
6			a term of
7	can	you give me	LOAN
8	could		
9	would		
10	can	I have	a term of
11	help	please	
12			

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

SYSTEM AND METHOD FOR CREATING A LANGUAGE GRAMMAR

Field of the Invention

5

This invention relates generally to computerized natural language systems. More particularly, it relates to a computer system and method for creating a grammar and automatically generating a corpus of valid utterances in the grammar, as well as a file containing the set of rules that define the grammar. It also relates to including token information in such a corpus to represent the meaning of each utterance to a speech-enabled system using the grammar.

15

Description of the Related Art

Computers have become a mainstay in our everyday lives. Many of us spend hours a day using the machines at work, home and even while shopping. Using a computer, however, has always been on the machine's terms. A mouse, pushbuttons and keyboards have always been somewhat of an unnatural way to tell the computers what we want. However, as computer technology continues to advance, the computer is edging towards communicating with humans on our terms: the spoken word.

There are essentially two steps in creating a computer that can speak with humans. First, the computer needs an automatic speech recognition system to detect the spoken words and convert them into some form of computer-

30

readable data, such as simple text. Second, the computer needs some way to analyze the computer-readable data and determine what those words, as they were used, meant.

5 Traditional speech recognizers have become quite efficient at identifying spoken words, and several good speech recognizers are commercially available. The ASR-1500, manufactured by Lernout & Hauspie; Watson 2.0, manufactured by AT&T; and Nuance 5.0, by Nuance are just a
10 few examples of effective continuous speech, speaker independent speech recognizers. A speech recognizer is "continuous" if it does not require the speaker to pause between words. A recognizer is "speaker independent" if it does not need to have heard the speaker's voice before
15 in order to understand the speaker's words. For example, some speech recognizers must first learn the user's voice before it can understand the words. Learning the user's voice usually means having the recognizer record and memorize the user's voice as s/he recites a list of key
20 phonetic words or sounds.

Recognizers, such as the ASR-1500 and Watson 2.0 listed above, typically require advance notice of the grammar they will be asked to listen for. This advance
25 notice comes in the form of a vendor-specific ASR grammar file that describes the rules and contents of the grammar. However, there is no easy way to create this data file. The manufacturers of the speech recognizers listed above will provide the user with the format in which the data
30 file must be written, but it is up to the user to actually write the data file.

There are several challenges to the user who must write this file: 1) the file format is usually what is best for the speech recognizer software, and as such is not very intuitive or easy to understand; 2) even simple grammars can result in large grammar files, increasing the possibility for error; 3) moderately large grammar files can have hundreds (or thousands) of rules, and logical errors in these rules are easy to make but hard to find, and 4) the large amount of time necessary to write the file and check for errors. What is needed is a simpler way for the user of a speech recognizer to define a grammar and generate a grammar file compatible with the speech recognizer.

Furthermore, in the related copending application noted above, a speech enabling system is disclosed which requires an annotated corpus for the grammar. The annotated corpus is essentially a listing of valid utterances (statements) in the grammar, as well as a token data value that represents the meaning for each valid utterance. Such a corpus could contain millions of entries, and what is needed is a simplified method for generating this annotated corpus.

The present invention satisfies both of these needs.

30

Summary of the Invention

A general purpose of the present invention is to provide a toolkit for generating a context-free grammar.

5 Another object of the present invention is to provide a toolkit for simplifying the task of generating a context-free grammar and an annotated corpus of utterances in the grammar.

10 A further object of the present invention is to provide a system and method for creating a corpus supported by a grammar that includes token data representing the meaning of each item in the corpus.

15 These and other objects are accomplished by the present invention which provides interactive voice recognition system developers with a simple and intuitive way to generate grammar corpora, or the list of words and phrases (utterances) that make up a language. A table
20 comprised of columns and rows of cells is initially used to define the grammar.

 As a simplified description of the tables used in the preferred embodiment, a valid utterance begins with
25 a cell in the leftmost column, ends in a column to the right, and must include the contents of one cell from each column in between. This left-to-right progression traces a path through the table, where the path itself is the valid utterance. Since it is possible for valid
30 utterances with completely different meanings to begin with the same cell in the leftmost column, heavy lines within columns are used to form barriers which the path

cannot cross when the next cell in the path is considered. Thus, the path of an utterance having one meaning cannot continue with a cell in the path of an utterance having a completely different meaning.

5

Eventually, the path of a valid utterance ends when the last word in the utterance is reached. In the preferred embodiment, the next non-blank cell in the same row as the last word of the utterance will contain token data. This token data will represent, to the speech-enabled application being developed, the meaning of the valid utterance.

Cells can also contain other types of data.

15 Some cells may contain references to other tables. Other cells may contain variables, which are phrases (such as numbers, dates, times and dollar amounts) that return a separate data value to the speech-enabled application. Cells may contain directives, or commands, which are

20 understood by the speech recognizer. Cells may be identified as optional, where utterances are valid with or without the contents of the optional cell.

When the developer is finished entering data in

25 the table, the grammar has been defined. The table is then traversed to generate an enumerated list (corpus) of valid utterances in the grammar, where each utterance in the corpus also includes a token. The table is also processed a second time to generate a grammar file which

30 will serve as input to the automatic speech recognition (ASR) system. During this second process, the grammar is also optimized to take advantage of performance

peculiarities of the ASR speech recognizer. These processes need not occur in the order in which they are described.

5 The corpus and grammar file are then used, in the preferred embodiment, in the interactive voice recognition system (IVR), which is the system that is actually conversing with the speaker. Valid utterances heard by the speech recognizer are reported to the IVR,
10 which then passes the valid utterances to the runtime interpreter (subject of copending application named above), which compares the valid utterances detected by the speech recognizer with the corpus of valid utterances in the grammar. When a match is found, the associated
15 token is retrieved and returned to the IVR. The IVR will understand, once given the token, what the detected valid utterance meant and can react accordingly.

Brief Description of the Drawings

20

Figure 1 shows an overview of an embedded natural language understanding system.

Figure 2 shows the grammar development toolkit main screen.

25

Figure 3 shows a tree representation of the grammar table in figure 2.

Figure 4 shows the various cell dialogs.

Figure 5 is a table showing the variable types supported in the preferred embodiment.

30

Figure 6 is a flow diagram showing the use of the grammar development toolkit.

Figure 7 shows the Annotated Corpus Dialog.

Figures 8a and 8b show sample formats for the annotated ASR corpus file and vendor-specific ASR grammar file, respectively.

5 Figure 9 is a flow diagram showing the operation of the IVR as it accesses the runtime interpreter.

Figures 10a-d are flow diagrams showing the steps taken by the toolkit to extract grammar rules from a table.

10

Description of the Preferred Embodiment

Before describing the present invention, several terms need to be defined. These terms, and their
15 definitions, include:

annotated ASR corpus file - data file containing a listing of valid utterances in a grammar, as well as token data for each valid utterance which represents the meaning of the valid utterance to the interactive voice
20 recognition system (IVR 130).

automatic speech recognition (ASR) - generic term for computer hardware and software that are capable of identifying spoken words and reporting them in a computer-readable format, such as text (characters).

25 cells - discrete elements within the table (the table is made up of rows and columns of cells). In the example rule given with the definition of 'rule' below, each of "I want", "I need" and "food" would be placed in a cell. Furthermore, in the preferred embodiment, the cells
30 containing "I want" and "I need" are vertically adjacent to one another (same column). Vertically adjacent cells

are generally OR'd together. The cell containing "food", however, would occur in the column to the right of the "I want" and "I need" column, indicating the fact that "food" must follow either "I want" or "I need" and as such, the
5 cell containing "food" will be AND'd to follow the cells containing "I want" and "I need".

constrained grammar - a grammar that does not include each and every possible statement in the speaker's language; limits the range of acceptable statements.

10 corpus - a large list.

grammar - the entire language that is to be understood. Grammars can be expressed using a set of rules, or by listing each and every statement that is allowed within the grammar.

15 grammar development toolkit (104) - software used to create a grammar and the set of rules representing the grammar.

natural language understanding - identifying the meaning behind spoken statements that are spoken in a
20 normal manner.

phrase - the "building blocks" of the grammar, a phrase is a word, group of words, or variable that occupies an entire cell within the table.

rules - these define the logic of the grammar.
25 An example rule is: ("I want" | "I need")("food"), which defines a grammar that consists solely of statements that begin with "I want" OR "I need", AND are immediately followed with "food".

runtime interpreter (124) - software that
30 searches through the annotated corpus (122) whenever a valid utterance is heard, and returns a token representing the meaning of the valid utterance.

runtime interpreter application program
interface (RIAPI) - set of software functions that serve
as the interface through which the interactive voice
response system (130) uses the runtime interpreter.

5 speech recognizer (116) - combination of
hardware and software that is capable of detecting and
identifying spoken words.

 speech recognizer compiler (114) - software
included with a speech recognizer (116) that accepts, as
10 input, a vendor-specific ASR grammar file (112) and
processes the file (112) for use in a speech recognizer
(116) during runtime.

 table - two dimensional grid used to represent a
grammar. Contents of a table are read, in the preferred
15 embodiment, from left to right.

 token - each valid utterance in the table is
followed by a cell that contains a token, where the token
is a unique data value (created by the developer when s/he
develops the grammar) that will represent the meaning of
20 that valid utterance to the interactive voice response
system (130).

 utterance - a statement.

 utterance, spoken - an utterance that was said
aloud. The spoken utterance might also be a valid
25 utterance, if the spoken utterance follows the rules of
the grammar.

 utterance, valid - an utterance that is found
within the grammar. A valid utterance follows the rules
which define the grammar.

30 variable - "place holder" used in the corpus
(122) to represent a phrase which has too many
possibilities to fully enumerate. For example, the

utterance "My favorite number between one and a million is
xxx" could result in 999,998 corpus entries, one for each
possible number. In the present invention, however, a
variable would be used to represent the number in the
5 corpus (122). Thus, a reduced corpus (122) would include
just one entry for this utterance: "My favorite number
between one and a million is [INTEGER]". The runtime
interpreter (124) is able to identify this variable in the
corpus, and performs additional processing during runtime
10 to interpret the number.

vendor-specific ASR grammar file (112) - a data
file that contains the set of rules representing a
grammar, and is written in a format that will be
recognized by the speech recognizer compiler (114).

15

Referring now to the drawings, where elements
that appear in several drawings are given the same element
number throughout the drawings, the structures necessary
to implement a preferred embodiment of an embedded natural
20 language understanding system (100) are shown in figure 1.
The basic elements comprising:

an interactive voice response system (130), or
IVR;

the grammar development toolkit (104);

25 a compiler (114) and speech recognizer (116),
which are part of an automatic speech recognition (ASR)
system (118);

an annotated ASR corpus file (122);

a vendor-specific ASR grammar file (112);

30 the runtime interpreter (124); and

the custom processor interface (126), or CP; and

the runtime interpreter application program interface (128), or RIAPI. The grammar development toolkit (104) is the focus of this application and is discussed in detail further below. The other elements
5 listed above are discussed in detail in the copending application cited above. However, a general overview of the speech-enabled system will be helpful for a full understanding of the operation and purpose of the toolkit (104).

10

1. Overview of Embedded Architecture

The following overview discusses the embedded architecture, which employs a single runtime interpreter
15 (124) which may be embedded within the RIAPI (128). There is a second, distributed, architecture which employs a plurality of runtime interpreters. The distributed architecture is discussed in the copending application referenced above.

20

The first step in implementing a natural language system is creating the set of rules that govern the valid utterances in the grammar. As an example, a grammar for the reply to the question: "what do you want
25 for lunch?" might be represented as:

```
<reply>: (("I want"|"I'd like")("hotdogs"|"hamburgers"));
```

Under this set of rules, all valid replies consists of two parts: 1) either "I want" or "I'd like", followed by 2) either "hot dogs" or "hamburgers". This notation is
30 referred to as Backus-Naur-Form (BNF), where adjacent elements are logically AND'd together, and the '|'

represents a logical OR. The preferred embodiment of the present invention generates this type of grammar.

Referring to Figure 1, the grammar is generated
5 by a developer using the grammar development toolkit
(104).

The toolkit (104) is a novel spreadsheet-oriented software package that provides the developer of a natural language application with a simplified way of generating a grammar.
10 In the preferred embodiment, the toolkit (104) resides on a computer that contains a central processing unit (102), application specific software (106), memory files (108) and input device such as keyboard (110).

15 When the developer has completed the grammar using the toolkit (104), two outputs are generated by the toolkit (104) for use in the natural language system. The first such output is a vendor-specific ASR grammar file (112), which is saved in a format that will be
20 recognizable by the speech recognizer (116). Speech recognizer (116) is a continuous speech, speaker independent speech recognizer. Commercially available speech recognizers (116) include the ASR-1500, manufactured by Lernout & Hauspie; Watson 2.0,
25 manufactured by AT&T; and Nuance 5.0, by Nuance. The preferred embodiment of the toolkit (104) is able to generate grammar files for any of these recognizers.

The vendor-specific ASR grammar file (112)
30 contains information regarding the words and phrases that the speech recognizer (116) will be required to recognize, written in a form that is compatible with the recognizer.

The file is also optimized to take advantage of peculiarities relating to the chosen speech recognizer (116). For example, experience with the L&H recognizers has shown that L&H grammars perform well if the grammar
5 avoids having multiple rules with the same beginning (three rules starting with "I want"). Optimization of a grammar for an L&H recognizer would rewrite a set of rules from <rule1>:(ab)|(ac)|(ad), to <rule2>:a(b|c|d). Here the three rules of 'rule1' have been rewritten and
10 combined into the one rule of 'rule2'.

In order to operate and recognize speech, the speech recognizer will need to compile the vendor-specific ASR grammar file (112) using a compiler tool (114)
15 supplied by the recognizer vendor. The preferred embodiment of the toolkit (104) knows, when the grammar is first generated, which speech recognizer (116) will be used and is able to format the vendor-specific ASR grammar file (112) accordingly.

20 The second output from the toolkit (104) is an annotated ASR corpus (122), which is actually a pair of flat files. The first of the pair is a corpus file, and contains a listing of all possible logical sentences or
25 phrases in the grammar (with the exception of variables, discussed below), the compartments (groups of tables) in which they appear, and a value representing the class of the utterance (sentence) heard. The second is an answers
30 file that maps each utterance class with a token, or data value that represents the meaning of the utterance. These two files will be used by the runtime interpreter (124).

During runtime, a speaker speaks into the microphone (or telephone) (120) attached to the speech recognizer (116). The recognizer (116) identifies the words and phrases it hears and notifies the IVR (130) when
5 a valid utterance has been heard. The IVR (130) is the system which needs the speech understanding capabilities, and includes the necessary external connections and hardware to function (for example, a banking IVR - 130 might include a connection to the bank database, a keypad
10 for entering data, a visual display for displaying information, a dispenser for dispensing money, and a speaker for speaking back to the user). This valid utterance is passed, in a computer-readable form such as text, to the IVR (130) which then notifies the runtime
15 interpreter (124) of the utterance that was heard. The runtime interpreter (124) consults the annotated ASR corpus (122) and returns an appropriate token to the IVR (130) for the valid sentence heard by the recognizer (116). This token represents the meaning of the utterance
20 that was heard by the recognizer (116), and the IVR (130) is then able to properly respond to the utterance. The CP (126) and RI-API (128) serve as software interfaces through which the IVR (130) may access the runtime interpreter (124). It is the IVR (130) that ultimately uses the
25 speech capabilities to interact with the speaker during runtime.

2. The Grammar Development Toolkit

30 In the preferred embodiment, the toolkit is developed using "Visual Basic" (trademark Microsoft Corp.) operating on a "Pentium"-based (trademark Intel Corp.)

computer system. The toolkit is designed to run on a "Windows NT" (trademark Microsoft Corporation). However, it is understood that the present invention can be developed and run on other computer systems using
5 different software.

Simply put, the toolkit is a software tool which enables a developer to visualize and create optimized ASR grammars for specified speech recognizers that use a BNF-
10 based language model for recognizing speech. One novel aspect of the toolkit is the spreadsheet format used to visualize the grammar being created. Figure 2 shows a typical main toolkit screen.

15 In the preferred embodiment, the toolkit main screen (fig. 2) displays a table (206) that is within a project. A project is defined as the set of compartments and tables which form the basis of a particular speech application. Compartments contain the one (or more)
20 table(s) related to a particular task within a project. A table essentially contains the tree for valid utterances. Thus, a typical project would be "banking". The "banking" project would contain compartments for the various tasks related to banking such as "account inquiry" and "loan
25 inquiry". Tables within the "account inquiry" compartment would then define the valid utterances which pertain to the topic of an account inquiry.

A table (206) is a matrix where the developer
30 enters data which will be used to generate the grammar, and is composed of rows and columns of cells. Each cell can contain data of one of the following types: Terminal

symbols, Non-terminal symbols, Variables and ASR Directives. Terminal symbols are any ASCII string (with the exception of special characters that have specific meaning to either the runtime interpreter 124 or the
5 speech recognizer 116), and they are the basic words and phrases of the language written in text form. Special characters in the preferred embodiment include '!', '/', '*' and '#' for the L&H recognizer, and parentheses, brackets, braces and carets in general.

10

Non-terminal symbols serve as a cross-reference to other tables. For example, "LOAN" in row 1, col. 3 and row 7, col. 3 of main table 206 is a non-terminal which references the sub-table 205a. When the main table is
15 processed to generate the corpus file (122) mentioned above, Non-terminal symbols will be replaced with the referenced table. "LOAN" will be replaced with the contents of sub-table 205a.

20

Variables are a type of Non-terminal symbol that, in addition to referencing another table, return a value to the controlling software (i.e., IVR) during runtime. Col. 3, row 4 in main table 206 is a cell containing the variable "YEARS" which references sub-table
25 205b. The table referenced by a variable defines the grammar for the variable and will be processed as any other referenced table when the vendor-specific ASR grammar (112) is created. However, variables are not completely enumerated when the corpus file (122) is
30 generated. This is because there would be too many possibilities to list if efficiency is to be maintained. For example, the valid utterance "My favorite number from

1 to a thousand is xxx" results in a thousand different possibilities. Instead of enumerating all of these possibilities, utterances that include variables are written to the corpus file (122) with special characters
5 "holding the place" of where variables will occur in the utterance. Thus, our example utterance would have a single corpus file (122) entry: "My favorite number from 1 to a thousand is [INTEGER1]". When the runtime interpreter (124) searches the corpus file (122) for a
10 detected utterance that includes variables, only the non-variable portion of the detected utterance must match. The variable portion will be interpreted separately using algorithms located in the runtime interpreter (124) (in the example, the algorithm for INTEGER1) and resulting
15 values will be stored in system memory for the IVR (130) to retrieve. In its present embodiment, the variable types which are supported with such algorithms are shown in figure 5.

20 In the corpus file (122), variables are set off from the regular text with square brackets ('[' and ']'). The characters between the brackets identify the variable type and the algorithm which will be processed at runtime when the variable is heard.

25 ASR Directives are special commands meant for the speech recognizer (116). These cells will appear in the grammar file (112), but not in the corpus file (122). For example, ASR Directives are used in the grammar file
30 (112) to inform the recognizer that there is a variable in the utterance and another table should be referenced.

Cells can also be Marked, Underlined, or characterized as Optional. These features will be explained during the discussion of table traversal below.

5 All valid utterances in the grammar represented by the table can be obtained by traversing the table from left to right. A valid utterance begins when one of the phrases in the first (leftmost) column is heard. When this first phrase is heard, the range of row numbers for
10 the phrases that are allowed to follow is determined by going up from the first phrase until a barrier line is encountered that also appears in the next column to the right, and going down until a barrier line is encountered which also appears in the next column to the right.
15 Barrier lines are created by underlining cells. These barrier lines form a barrier in the table which is not to be crossed during the enumeration traversal below. The next phrase must be found (if the sentence is to be valid) in the first column to the right of the current column
20 that contains entries within this range of rows. This process is continued until the next non-blank column is a marked column. Each compartment has a main table, and the rightmost column of this main table is the marked column which contains the token to be returned to the IVR
25 (130) for each valid utterance.

For example, using the table in fig. 2 (and the tree in figure 3), if the phrase "I think" was first heard, the range of row values for the next valid phrase
30 is 1 (got to top of column without encountering a barrier line) to 6 (can't cross barrier line under row 7 because it also appears in column 2). The next column that

contains entries within this range of rows (skipping any blank ones) is column 2, so the next valid phrase must be found in column 2 between rows 1 and 6. If the phrase "I want" was heard next, it would be valid because this

5 phrase exists in row 2 (between 1 and 6) of column 2. Once "I want" is heard, the process of determining the next valid range begins again, but now starts from the cell containing "I want". Going up and down as before, the next valid range of rows also goes from row 1 to 6,

10 but the next column with entries in this range is column 3. So, the next valid phrase must occur in column 3, between rows 1 to 6. If "the term to be" is heard next, it would be valid. Then, the next valid range of rows would go from 4 to 6 (due to the heavy lines under rows 3

15 and 6 in the column containing "the term to be"), and the next column with entries in this range of rows is column 4. The traversal is now looking for the "YEARS" variable (20, 30, or 60), and this continues until a complete utterance is processed and the marked column (column 7 -

20 "TOKEN") is reached.

During the traversal discussed above, a second "mini-traversal" will take place whenever a non-terminal cell is encountered. A non-terminal cell, as discussed

25 above, simply identifies another table which should appear in place of the non-terminal cell. When a non-terminal cell is encountered during a traversal, the referenced table is traversed in the same manner, and any of the utterances from this "sub-table" may be placed in the non-

30 terminal cell within the current utterance of the main traversal. Functionally, it is as if the referenced table replaced the non-terminal cell in the current table.

Cells may also be characterized as Optional. Each utterance that includes an optional cell will be valid with or without the contents of that cell. Thus, 5 when such an utterance is processed into the annotated ASR corpus (122), the optional cells in the utterance result in multiple corpus items (utterances) to account for the various possibilities. ASR Directives are used in the vendor-specific grammar file (112) to inform the speech 10 recognizer (116) that a particular cell is optional.

Finally, it should be noted that the traversal discussed above demonstrates how a table can be used to check the validity of a given utterance. Another type of 15 traversal is actually performed by the toolkit (104) software. This second type of traversal is used to enumerate all possible valid utterances from the table (with a few exceptions, such as variables), or in other words, the enumeration traversal seeks to "see what's at 20 the end of every valid path through the table" and record the valid utterance formed by each path.

In the preferred embodiment, when the enumeration traversal takes place, the software logically 25 travels completely down one path until it ends (in a Marked column), recording the contents of the cells that it passed through. At that point, one valid path has been completed and the contents of the cells that the path passed through are written, in the order they were 30 encountered, as a valid utterance. The traversal then backs up until the last "branch" (where more than one cell is valid) in the path, maintains the recording of the

cells in the path leading up to the branch, and goes down the other "branch" until the end, backs up . . . and so on until all possible paths have been enumerated.

5 Using the table in figure 2 again as an example, an enumeration traversal might start with the valid utterance "I think I need a mortgage with a twenty year term" (where 'twenty' would be a YEAR variable). The next
10 valid utterance, after "backing up" from the end of the first valid utterance until a branch, would be "I think I need a mortgage that has a twenty year term". The branch occurred in column 4, where "with a" and "that has a" both would have resulted in valid utterances. This is merely the type of enumeration traversal performed in the
15 preferred embodiment of the invention, and it is understood that various methods of traversal would also result in a full enumeration. In the preferred embodiment, the list of valid utterances in the corpus (122) need not be in any particular order, so any
20 enumeration traversal method would work. As a note, "thirty" and "sixty" are not considered branches in the preferred embodiment because they are all part of the single variable YEAR, and variables are not completely enumerated. Also, since "I think" is in an optional cell,
25 the valid utterances enumerated above would also be valid if "I think" were removed, so a second valid utterance will be enumerated as well.

30 A tool (202) is displayed across the top of the toolkit main screen in figure 2. This tool allows the user to manage compartments and projects. Under the 'File' menu, the user can create, open or close projects,

NOT FURNISHED UPON FILING

Variable" to make a cell a variable; "Insert Subgrammar" to specify a non-terminal and insert the table (or subgrammar) referenced by the non-terminal; and "Insert ASR Option" to insert an ASR directive.

5

When a cell's attributes are set to Variable, Non-terminal or ASR Directive, one of the dialogs in figures 4a-c will appear. Figure 4a shows the Non-terminal specifications dialog (400). As defined above, a
10 Non-terminal cell type references another table. This dialog (400) simply prompts the developer for the name of the referenced table.

15 Figure 4b shows the ASR Directives Dialog (410), which provides the user with a listbox of the available ASR Directives for the current speech recognizer (116). If the selected ASR Directive requires additional argument data, the developer can provide this additional argument
20 data in the dialog.

Figure 4c shows the Variable Specification Dialog (420), which is displayed when the developer designates a variable cell. The dialog (420) prompts the
25 developer for the name of the variable, the table that is referenced by the variable, and the variable type. A list of variable types supported by the preferred embodiment, as well as the value type to be returned by the variable, is found in figure 5.

30

Figure 6 shows the steps a developer will take when using the toolkit (104) to create a grammar. The

developer begins in step 600 by opening the toolkit (104) program. In a typical windowing environment, this will involve selecting an icon. While the preferred embodiment of the toolkit (104) is as an OLE server for inclusion in another vendor-developed toolkit, the toolkit could also
5 be made in a standalone version.

Once the program has begun, the developer will need to open a new project to work on in step 602. The
10 developer then creates a compartment in step 604, and defines the number of columns and rows for the compartment's main table in step 606. This can be done using the sizing buttons (204b). When the size of the compartment main table has been established, the next step
15 (608) is to fill in the cells that are to contain terminal data.

As discussed above, terminal data is simply text data that defines the non-variable phrases ("I'd like", "I want", etc.) which are valid in the grammar.
20

In step 610, the developer fills in the variable cells (if any) of the table. In the preferred embodiment, the variables dialog shown in figure 4c is used to define the variable, which includes the variable name and a
25 reference to the algorithm which will be performed when the variable is processed (the actual algorithm is located within the runtime interpreter 124).

In step 612, the developer fills in the cells
30 (if any) which are to contain ASR directives. In the preferred embodiment, the dialog in figure 4b is used to select the ASR directive from a list of supported ASR

directives for the current speech recognizer (116) (the current speech recognizer is chosen from a list in the 'File' menu of the tool - 202).

5 In step 614, the developer identifies which cells (if any) are to be optional by using the right mouse button and the "Make Optional" option of the menu shown in figure 2b.

10 In step 616, the developer identifies and fills in the column in the compartment main table that contains the token data. In the preferred embodiment, the rightmost column of the compartment main table is marked by default as having token data.

15 In step 618, the developer identifies the barriers used during the traversal discussed above. In the preferred embodiment, this is done by underlining certain cells using the menu shown in figure 2b.

20 In step 620, the developer fills in the non-terminal cells. In the preferred embodiment, this is done using the dialog shown in figure 4a. The developer will also need to create a new table to contain the subgrammar, 25 which can be done using the 'table' menu. This table will also need the terminal, non-terminal, variable, optional and ASR directive cell steps discussed above if the referenced table requires any of those cell types. However, in the preferred embodiment, referenced tables do 30 not have marked columns and do not return tokens.

In step 622, the compartment is complete. If another compartment is needed for the project, the developer proceeds to step 604 and creates the new compartment. If no other compartments are necessary, the developer can proceed to either of two steps, 624 and 626. In step 624, the grammar file (112) is generated for the ASR (114, 116). In the preferred embodiment, this process is initiated using the 'File' menu option in the compartment window (203). This generation is done by performing an analysis of the table, and results in a file (112) stored in memory. This file (112) contains the grammar rules embodied in the tables and is in a form that can be compiled and used by the speech recognizer (116).

The analysis performed during generation of the grammar file (112) seeks to generate a set of BNF rules which are followed by valid utterances in the grammar. The generation of the BNF rules begins with a table that has no non-terminals. All non-terminals are replaced with the tables that they reference, resulting in a complete, expanded main table.

Consider the following Tables 1a-e.

TABLE 1a:

Column	1	2	3	4
Row 1	A	B	C	D
2	E	F	G	H
3	I	J	K	L
4	M	N	O	P

Rule: "(A|E|I|M) (B|F|J|N) (C|G|K|O) (D|H|L|P)".

TABLE 1b:

Column	1	2	3	4
Row 1	A	B	C	D
5 2	E	F	<u>G</u>	H
3	I	<u>J</u>	K	L
4	M	N	O	P

Rule: "(A|E|I|M)((B|F|J)|N)((C|G)(D|H))|((K|O)(L|P))"

10 TABLE 1c: (with group number)

Column	1	2	3	4
Row 1	A(1)	B(2)	C(2)	D(2)
2	E(1)	<u>F(2)</u>	<u>G(2)</u>	<u>H(2)</u>
3	<u>I(1)</u>	<u>J(3)</u>	K(4)	L(4)
15 4	M(5)	N(5)	O(4)	P(4)

Rule: ((A|E|I)((B|F)(C|G)(D|H)))|(J(K|O)(L|P))|(MN((K|O)(L|P)))

TABLE 1d:

Column	1	2	3	4
Row 1	<u>A</u>	<u>B</u>	C	D
2	E	F	G	H
3	I	J	<u>K</u>	<u>L</u>
4	M	N	O	P

25 Rule: "((AB)|((E|I|M)(F|J|N)))((C|G|K)(D|H|L))|(OP)"

The main process, in the preferred embodiment, of generating the grammar rules which represent the table is shown in the flow diagram of Figures 10a-d. Titled
 30 "HandleRegion", this flow diagram shows the recursive algorithm used. HandleRegion itself receives a table, or portion of a table, as input. In the preferred

embodiment, any non-terminals originally found in the tables given to HandleRegion are replaced with the sub-tables that they reference before HandleRegion is called. Tables are represented by a two-dimensional array, and
5 defined by their corners (i.e. row1,col1 - row4,col4). HandleRegion assembles the grammar rules defined by the table by recursively dividing the table into smaller and smaller tables ("regions") until "simple" regions are formed. A simple region is a region of cells in which
10 there are no heavy lines (barriers). Table 1a depicts a simple region.

Once HandleRegion is begun (step 1002), a check is made to determine whether the region to be handled is a
15 simple one. If a rule is to be extracted from table 1a, then this check succeeds. When the region is a simple one, vertically adjacent cells are grouped using logical 'OR' (step 1006). In table 1a, four groups would be formed, one for each column. The groups would be,
20 (A|E|I|M), (B|F|J|N), (C|G|K|O) and (D|H|L|P), where the '|' represents logical 'OR'. Then, these groups would be combined in a left to right order using logical 'AND' (step 1008), resulting in
"(A|E|I|M)(B|F|J|N)(C|G|K|O)(D|H|L|P)". This rule is
25 returned in step 1010 and HandleRegion ends in step 1072.

If the region to be handled is not a simple one, such as the one shown in table 1b, then the check in step 1004 fails, and the process moves on to step 1012. In
30 step 1012, the region to be handled is checked to see if it is a "prefixed" region. A prefixed region has no barrier lines in the first 'n' columns (these are simple

columns), and thus, table 1b is a prefixed region because it has no barrier lines in its first 1 column. In step 1014, the prefixed region is divided into two sub-regions. The first sub-region is the leftmost n columns (contains
5 A, E, I and M), and the second sub-region is the rest of the prefixed region. In step 1016, the rule for the prefixed region is defined as the rule for the first region AND the rule for the second region. HandleRegion is then recursively called once to find the rule for the
10 first region, and again to find the rule for the second region. In step 1018, this rule is returned and HandleRegion ends in step 1072.

If, in step 1012, the region is not a prefixed
15 region, it is checked (step 1020) to see if it is a "postfixed" region. A postfix region is the reverse of a prefixed region, and has no barrier lines in the last 'n' columns. If the region is a postfix region, it is divided in step 1022 into two sub-regions, one for the
20 rightmost n columns, and another for the rest. Step 1024 defines the rule for the region as the logical AND of the two sub-regions, much like step 1016 for the prefixed region. Again, the rule is returned (step 1026) and HandleRegion ends in step 1072.

25
If, in step 1020, the region is not a postfix region, it is checked (in step 1028) to see if there are any "overlapping" barriers in the region. Overlapping barriers occur in columns which have more than one
30 barrier, and table 1c contains an overlapped column (column 2). If overlapped columns exist in the region, then a cell-by-cell approach is used to determine the

simple regions contained within. In step 1030, the process starts with the top, left cell in the region to be handled. The contents of this cell ('A') are added to the current simple region being defined.

5

Then, in step 1032, the process attempts to go down the current column until a barrier line is encountered. Cells that the process enters are added to the current simple region. In table 1c, the process can go down until cell 'I', and each cell that it enters is added to the current simple region (so far, 'A', 'E' and 'I' are in the current simple region). At this time, a "ceiling" and "ground" are noted. The "ceiling" is the first barrier or border in the current column above the current cell (border above 'A'), and the "ground" is the first barrier or border in the current column that is below the current cell (barrier under 'I' in row 3). In step 1034, the process attempts to go right until either the "ceiling" or "ground" of the next cell to the right is different. In the example using table 1c, after 1032, the process is at cell 'I'. It then attempts to go right, to cell 'J'. The "ground" under 'J' is the same (barrier still under row 3), but the "ceiling" is different. The first barrier or border above 'J' is NOT the border above 'A' (which was the case for 'I'), but the barrier under 'F'.

Thus, the process has gone right until the "ceiling" changed. The process would also stop going right if the rightmost border of the region was encountered or if the cell to the right has already been assigned to a simple region by the overlapping process.

NOT FURNISHED UPON FILING

In step 1042, the overlapping process checks to see if all cells have been assigned to a simple region. If not, the process goes to step 1044, which chooses the next cell to be assigned to a simple region. This selection is done from left to right across the current row, and when a row is done, the first cell of the next row is chosen. In our example, the first simple region began with the top, left cell ('A'). The next cell to the right that isn't already part of a simple region is 'B', so 'B' is the next starting point.

Going again through steps 1032-1040 starting at 'B', the process goes down until (and including) 'F', right until (and including) 'H', up until (and including) 'D', and left until (and including) 'C' before becoming trapped at 'C'. Thus, the second simple region contains 'B', 'C', 'D', 'F', 'G' and 'H'. Since the second simple region started with 'B', the process considers 'C' as its next starting cell. 'C' is already assigned, so are 'D', 'E', 'F', 'G', 'H' and 'I'. Thus, the next starting cell will be 'J'.

Going through steps 1032-1040 starting at 'J', the process can't go down because of the barrier, can't go right because the ground would change, can't go up because of the barrier, and can't go left because 'I' has already been assigned to a simple region. So 'J' is its own simple region.

Going through steps 1032-1040 starting at 'K', the process goes down until 'O', right until 'P', up until 'L' before it is trapped at 'L'. The fourth simple region

NOT FURNISHED UPON FILING

would be BEGIN->1->3->4->END. Finally, in step 1054, each path is converted to a rule by discarding the BEGIN and END, substituting logical 'AND' for the '->' and replacing the region numbers with the simple region rules generated
5 in step 1046. These rules, one for each path, are then combined using logical 'OR' to form the one rule for the overlapping region. This one rule is returned in step 1056, and the HandleRegion process ends in step 1072.

10 If, in step 1028, the region to be handled is not an overlapping region, the HandleRegion process moves to step 1058, where the longest barrier line in the region to be handled is determined. If there is a tie, one is arbitrarily chosen. In step 1060, this line is checked to
15 see if it extends from one edge of the region to the other. If it does not, the region to be handled is split vertically into two horizontally adjacent sub-regions. This split is done at an arbitrary edge of the longest barrier line. In the preferred embodiment, the left edge
20 is used, and if the left edge of the longest line is the left border of the region, the right edge is used. In table 1d, the barrier under 'A' and 'B' is arbitrarily chosen over the one under 'O' and 'P'. Furthermore, the split occurs on the right edge of this barrier, forming
25 two sub-regions. The first sub-region includes columns 1 and 2, while the second sub-region contains columns 3 and 4. In step 1064, the HandleRegion process is called once for each sub-region, and the returned rules for the sub-regions are combined using a logical 'AND'.

30

If, in step 1060, the longest barrier line did extend across the entire region, step 1066 simply divides

the region into two sub-regions based on the longest line. HandleRegion is then called once for each sub-region, and these regions are combined, in step 1068, using a logical 'OR'. Step 1070 returns whatever rule was created (either
5 step 1064 or 1068), and 1072 ends the HandleRegion process.

As the grammar files (112) are being built, they will also be optimized. During optimization, the grammar
10 rules represented by the tables may be modified to accommodate peculiarities of the chosen speech recognizer (116) and enhance performance. For example, combining multiple rules that begin or end with the same words into a single rule increases efficiency for some recognizers.
15 If two rules begin with the same words (i.e. AB and AC), then the two rules can be combined into a single rule with an additional logical 'OR' (i.e. A(B|C)). Likewise, two rules that end with the same words (i.e. AC and BC), would be combined into a single rule (i.e. (A|B)C).

20 In step 626, which can occur before or after step 624, the annotated ASR corpus file is generated. In the preferred embodiment, this is initiated with a "generate corpus" in the 'File' menu of the compartment
25 window (203). In the preferred embodiment, individual corpora are first generated and stored in memory for each compartment, and the individual corpora are merged by selecting this option in the tool (202) 'File' menu. When this option is selected, the Annotated Corpus Dialog shown
30 in figure 7 appears and the developer selects which corpora will be merged as well as the name of the newly merged corpus. Merging essentially appends the listings

of the various corpora into a single listing. The resulting corpus file (122) is stored in memory and will be used by the interpreter (124) during runtime to process utterances that are heard by the speech recognizer (116).

5

When a compartment corpus is generated, the enumeration traversal described above is performed for the tables within the compartment. In the preferred embodiment, valid utterances within a compartment begin and end in the compartment main table, but the valid utterance "paths" may pass through several tables referenced by non-terminals. Each valid path through the table is written (in text form) to the compartment corpus. Each valid path will also include token data (which serves as a token class identifier) found at the end of the valid path in the compartment's main table (the marked column). As the compartment corpus is generated, the logic of the various valid utterances is also checked to ensure that the resulting grammar makes sense (no null utterances, one token class per valid utterance, etc.).

20

The resulting corpus (122) is actually a pair of files in the preferred embodiment, and the toolkit (104) generates both of these files when the various compartment corpora are merged into a single corpus. The preferred format for these files is as follows:

25

corpus file:

{1}:comp1 - first utterance in comp1, token class 1

{1-1}: comp1 - second utterance in comp1, token class

30 1

{1-2}: comp1 - third utterance in comp1, token class

1

{2}: compl - first utterance in compl, token class 2
etc..

The first number indicates the class of the utterance.

The second number, if any, denotes a member of this class.

- 5 The text that immediately follows, "compl", identifies the compartment. The rest of the line is the actual utterance in text form.

answer file:

10 {1}:
years
{2}:
days

- 15 The first number is the utterance class, and the next line contains the token to be returned. In the example, utterances of class '1' will return the token 'years' to the IVR (130), and utterances of class '2' will return the token 'days'.

20

- After the corpora are merged, the resulting corpus file (122) is ready for use at runtime by the runtime interpreter (124). With the corpus file (122) and grammar files (112) created, the developer has completed
25 development with the grammar development toolkit (104). The toolkit is terminated in step 628 by either selecting the 'Exit' option from the File menu, or by clicking on the 'x' in the upper right-hand corner of the toolkit window.

30

In light of the above teachings, it is understood that variations are possible without departing

from the scope of the invention embodied in these teachings. Any examples provided as part of the inventors' preferred embodiment are presented by way of example only, and are not intended to limit the scope of
5 the invention. Rather, the scope of the invention should be determined using the claims below.

What is claimed is:

Claims

- 5 1. A method for creating a language grammar for use in an interactive voice response system, comprising the steps of:
- opening a main table, said main table comprising rows and columns of cells;
- 10 entering phrase data in said cells;
- marking one of said cells as ending a valid utterance in said language grammar;
- entering token data in a cell associated with said marked cell, where said token data represents the
- 15 meaning of said valid utterance ended by said marked cell;
- performing an enumeration traversal of said main table a first time to enumerate valid utterances found in said main table;
- writing results of said enumeration traversal to
- 20 a first output file;
- analyzing said main table to extract a set of rules which define the valid utterances found in said main table; and
- writing results of said analysis to a second
- 25 output file.
2. The method of claim 1, further comprising the step of creating a barrier in said main table that is not to be crossed during said enumeration traversal.

3. The method of claim 1, further comprising the step of identifying a cell as a non-terminal cell, where the contents of said non-terminal cell reference another table.

5

4. The method of claim 3, further comprising the step of identifying a non-terminal cell as a variable cell, where the contents of said variable cell also reference an algorithm which will be performed when said interactive voice response system detects a variable in a spoken utterance.

5. The method of claim 1, further comprising the step of identifying a cell as an optional cell, where valid utterances that include said optional cell are valid both with and without contents of said optional cell.

6. The method of claim 1, further comprising the step of identifying a cell as an automatic speech recognition (ASR) system directive cell, where the contents of said ASR directive cell have special meaning to an ASR system.

7. The method of claim 1, wherein said step of performing an enumeration traversal comprises the following steps:

- a) beginning with the topmost, leftmost non-blank cell in said main table;
- b) recording contents of said topmost, leftmost non-blank cell in an utterance buffer;
- c) moving to next non-blank cell to the right and appending contents of said cell to said utterance buffer;

d) repeating step (c) until marked column is reached;

e) writing contents of said utterance buffer to said first output file;

5 f) moving left to the previous non-blank cell whose contents are in said utterance buffer and erasing contents of said previous non-blank cell from utterance buffer;

10 g) repeating step (f) until current cell has vertically adjacent non-blank cell below said current cell with no heavy line between said current cell and said vertically adjacent non-blank cell;

15 h) moving down to said vertically adjacent non-blank cell and appending contents of said vertically adjacent non-blank cell to said utterance buffer; and

i) repeating steps (b)-(h) until entire main table has been traversed.

20 8. The method of claim 1, wherein said step of analyzing further comprises the steps of:

j) grouping contents of vertically adjacent cells with logical OR; and

k) combining said grouped contents which are horizontally adjacent with logical AND.

25

9. A computer system for creating a language grammar, said computer system comprising:

a main table of columns and rows of cells;

30 terminal cells in said main table, where phrase data is entered in said terminal cells and are combined to form valid utterances in said language grammar;

marked cells in said main table, where marked cells indicate the end of a valid utterance;

token data in a plurality of said cells, where each said valid utterance is associated with said token
5 data which represents the meaning of said valid utterance;

means for performing an enumeration traversal of said main table, where said enumeration traversal automatically generates a list of valid utterances in said language grammar; and

10 means for analyzing said main table and extracting a set of rules from the contents of said main table, where valid utterances in said language grammar follow said set of rules.

15 10. The system of claim 9, further comprising one or more barriers in said main table that are not to be crossed during said enumeration traversal.

20 11. The system of claim 9, further comprising a non-terminal cell, where the contents of said non-terminal cell reference another table.

12. The system of claim 11, further comprising a variable cell, where the contents of said variable cell reference
25 an algorithm which will be performed when said interactive voice response system detects a variable in a spoken utterance.

30 13. The system of claim 9, further comprising an optional cell, where valid utterances that include said optional cell are valid both with and without contents of said optional cell.

14. The system of claim 9, further comprising an automatic speech recognition (ASR) system directive cell, where the contents of said ASR directive cell have special
5 meaning to an ASR system.

15. A method for creating a language grammar for use in an interactive voice response system, comprising the steps of:

10 opening a main table, said main table comprising rows and columns of cells;

entering phrase data in said cells;

creating a barrier in said main table that is not to be crossed during an enumeration traversal;

15 identifying a cell as a non-terminal cell, where the contents of said non-terminal cell reference another table.

further identifying a non-terminal cell as a variable cell, where the contents of said variable cell
20 also reference an algorithm which will be performed when said interactive voice response system detects a variable in a spoken utterance;

identifying a cell as an optional cell, where valid utterances that include said optional cell are valid
25 both with and without contents of said optional cell;

identifying a cell as an automatic speech recognition (ASR) system directive cell, where the contents of said ASR directive cell have special meaning to an ASR system;

30 marking one of said cells as ending a valid utterance in said language grammar;

entering token data in a cell associated with said marked cell, where said token data represents the meaning of said valid utterance ended by said marked cell;

performing said enumeration traversal of said
5 main table a first time to enumerate valid utterances found in said main table;

writing results of said enumeration traversal to a first output file;

analyzing said main table to extract a set of
10 rules which define the valid utterances found in said main table; and

writing results of said analysis to a second output file.

15 16. The method of claim 15, wherein said step of performing an enumeration traversal comprises the following steps:

a) beginning with the topmost, leftmost non-blank cell in said main table;

20 b) recording contents of said topmost, leftmost non-blank cell in an utterance buffer;

c) moving to next non-blank cell to the right and appending contents of said cell to said utterance buffer;

25 d) repeating step (c) until marked column is reached;

e) writing contents of said utterance buffer to said first output file;

f) moving left to the previous non-blank cell
30 whose contents are in said utterance buffer and erasing contents of said previous non-blank cell from utterance buffer;

g) repeating step (f) until current cell has vertically adjacent non-blank cell below said current cell with no heavy line between said current cell and said vertically adjacent non-blank cell;

5 h) moving down to said vertically adjacent non-blank cell and appending contents of said vertically adjacent non-blank cell to said utterance buffer; and

 i) repeating steps (b)-(h) until entire main table has been traversed.

10

17. The method of claim 15, wherein said step of analyzing further comprises the steps of:

 j) grouping contents of vertically adjacent cells with logical OR; and

15 k) combining said grouped contents which are horizontally adjacent with logical AND.

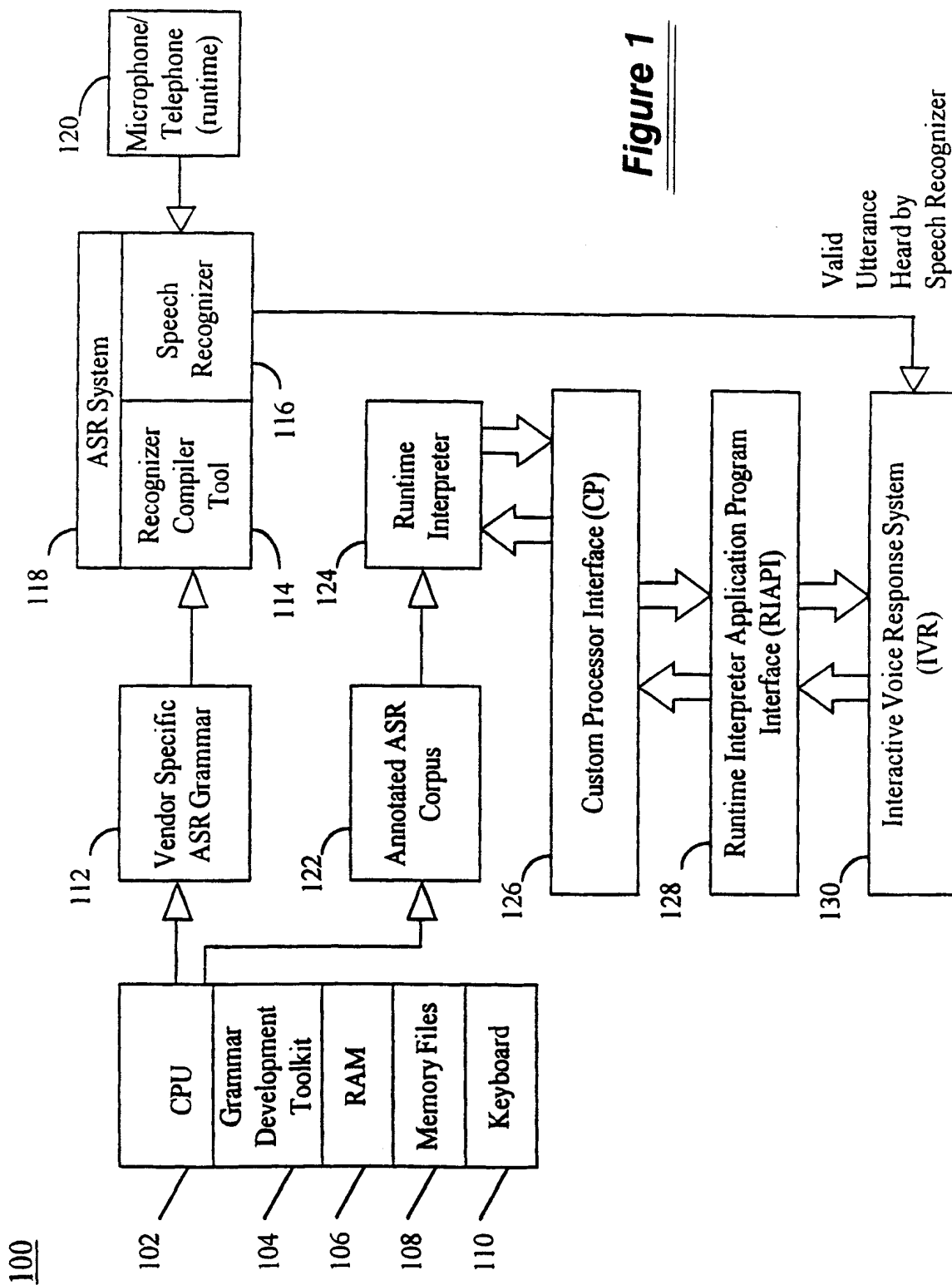
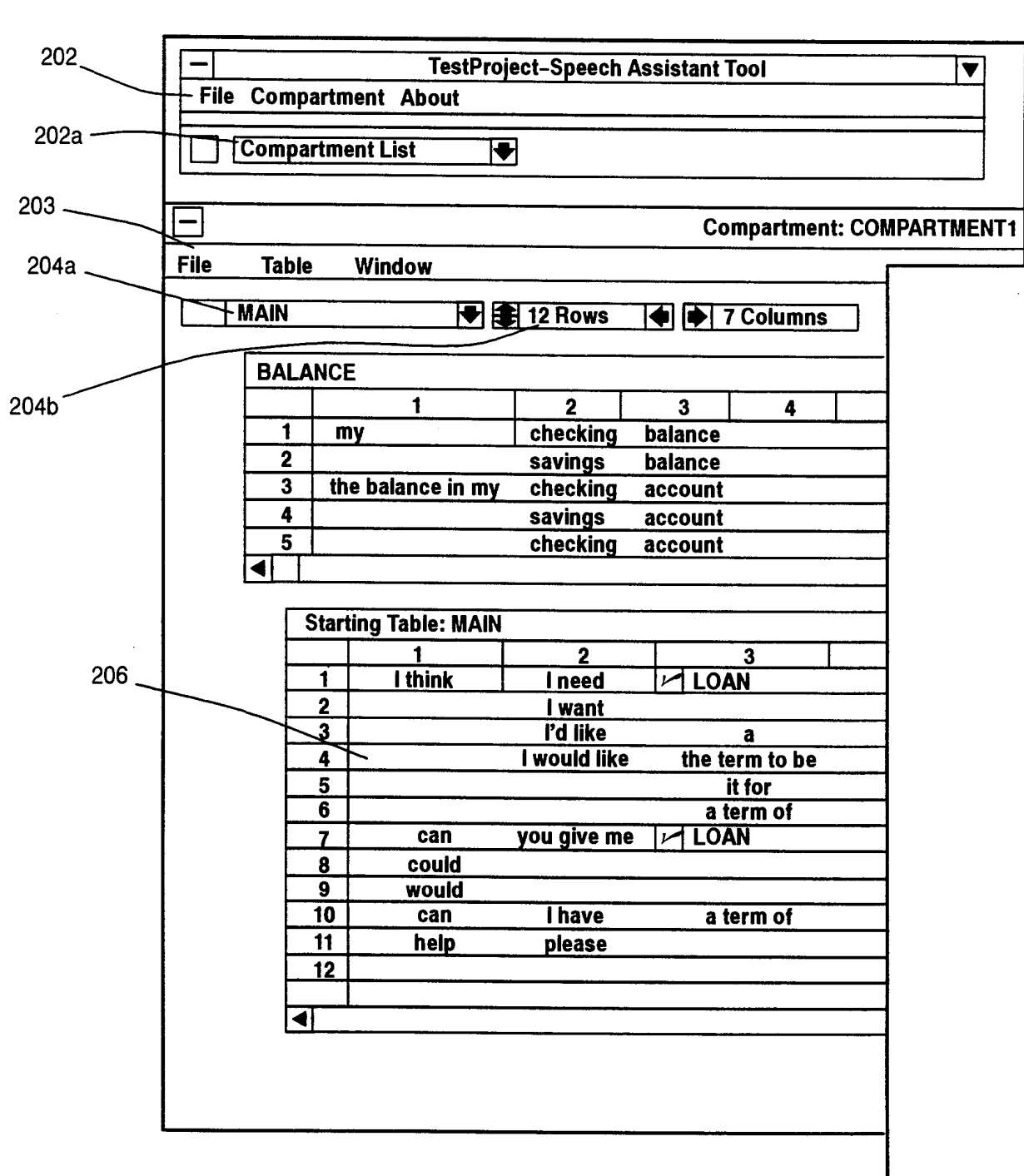


Figure 1

2 / 21

**Figure 2****Fig. 2****Fig. 2**

NOT FURNISHED UPON FILING

4 / 21

<div style="display: flex; justify-content: space-between; align-items: center;"> — TestProject-Speech Assistant Tool ▼ </div>																																																									
<div style="display: flex; justify-content: space-between; align-items: center;"> File Compartment About </div>																																																									
<div style="display: flex; justify-content: space-between; align-items: center;"> <input type="checkbox"/> Compartment List ▼ </div>																																																									
<div style="display: flex; justify-content: space-between; align-items: center;"> — Compartment: COMPARTMENT1 </div>																																																									
<div style="display: flex; justify-content: space-between; align-items: center;"> File Table Window </div>																																																									
<div style="display: flex; justify-content: space-between; align-items: center;"> MAIN ▼ 12 Rows ◀▶ 7 Columns </div>																																																									
<table border="1" style="width: 100%; border-collapse: collapse;"> <tr> <th colspan="5">BALANCE</th> </tr> <tr> <th></th> <th>1</th> <th>2</th> <th>3</th> <th>4</th> </tr> <tr> <td>1</td> <td>my</td> <td>checking</td> <td colspan="2">balance</td> </tr> <tr> <td>2</td> <td></td> <td>savings</td> <td colspan="2">balance</td> </tr> <tr> <td>3</td> <td>the balance in my</td> <td>checking</td> <td colspan="2">account</td> </tr> <tr> <td>4</td> <td></td> <td>savings</td> <td colspan="2">account</td> </tr> <tr> <td>5</td> <td></td> <td>checking</td> <td colspan="2">account</td> </tr> </table>		BALANCE						1	2	3	4	1	my	checking	balance		2		savings	balance		3	the balance in my	checking	account		4		savings	account		5		checking	account																						
BALANCE																																																									
	1	2	3	4																																																					
1	my	checking	balance																																																						
2		savings	balance																																																						
3	the balance in my	checking	account																																																						
4		savings	account																																																						
5		checking	account																																																						
<table border="1" style="width: 100%; border-collapse: collapse;"> <tr> <th colspan="4">Starting Table: MAIN</th> </tr> <tr> <th></th> <th>1</th> <th>2</th> <th>3</th> </tr> <tr> <td>1</td> <td>I think</td> <td>I need</td> <td><input checked="" type="checkbox"/> LOAN</td> </tr> <tr> <td>2</td> <td colspan="3">I want</td> </tr> <tr> <td>3</td> <td colspan="3">I'd like a</td> </tr> <tr> <td>4</td> <td colspan="3">I would like the term to be</td> </tr> <tr> <td>5</td> <td colspan="3">it for</td> </tr> <tr> <td>6</td> <td colspan="3">a term of</td> </tr> <tr> <td>7</td> <td>can</td> <td>you give me</td> <td><input checked="" type="checkbox"/> LOAN</td> </tr> <tr> <td>8</td> <td colspan="3">could</td> </tr> <tr> <td>9</td> <td colspan="3">would</td> </tr> <tr> <td>10</td> <td>can</td> <td>I have</td> <td>a term of</td> </tr> <tr> <td>11</td> <td>help</td> <td colspan="2">please</td> </tr> <tr> <td>12</td> <td colspan="3"></td> </tr> </table>		Starting Table: MAIN					1	2	3	1	I think	I need	<input checked="" type="checkbox"/> LOAN	2	I want			3	I'd like a			4	I would like the term to be			5	it for			6	a term of			7	can	you give me	<input checked="" type="checkbox"/> LOAN	8	could			9	would			10	can	I have	a term of	11	help	please		12			
Starting Table: MAIN																																																									
	1	2	3																																																						
1	I think	I need	<input checked="" type="checkbox"/> LOAN																																																						
2	I want																																																								
3	I'd like a																																																								
4	I would like the term to be																																																								
5	it for																																																								
6	a term of																																																								
7	can	you give me	<input checked="" type="checkbox"/> LOAN																																																						
8	could																																																								
9	would																																																								
10	can	I have	a term of																																																						
11	help	please																																																							
12																																																									

Figure 2b

Fig. 2b

Fig. 2b

▼▲

Hints

5	6	▲
ch		
sa		
ch		
sa		
ch		▼
		▶

		▲
1	20	
2	30	
3	60	▼
◀		▶

		▲
1	a mortgage	
2	a loan	
3		▼
◀		▶

Cut

Copy

Paste

Clear

Replace...

Underline

Make Optional

Insert Variable...

Insert SubGrammar...

Insert ASR Options...

Edit...

Print...

5	6	TOKEN	▲
years, YEARS]	year term	Myyears	
ar term			
	years		
years, YEARS]			
years		Tyears	
		help	
			▼
			▶

Figure 2b

Fig. 2b	Fig. 2b
---------	---------

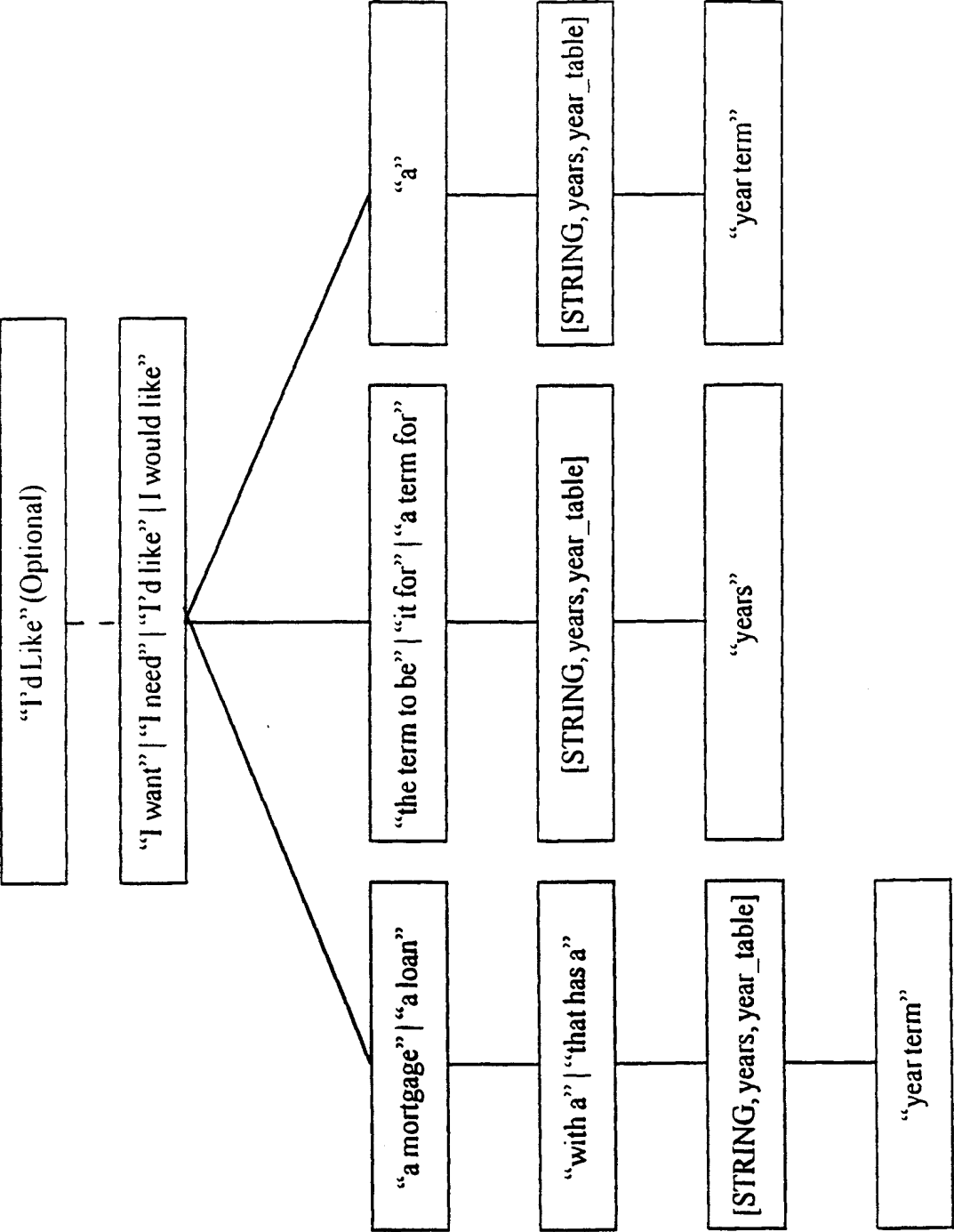


Figure 3

7/21

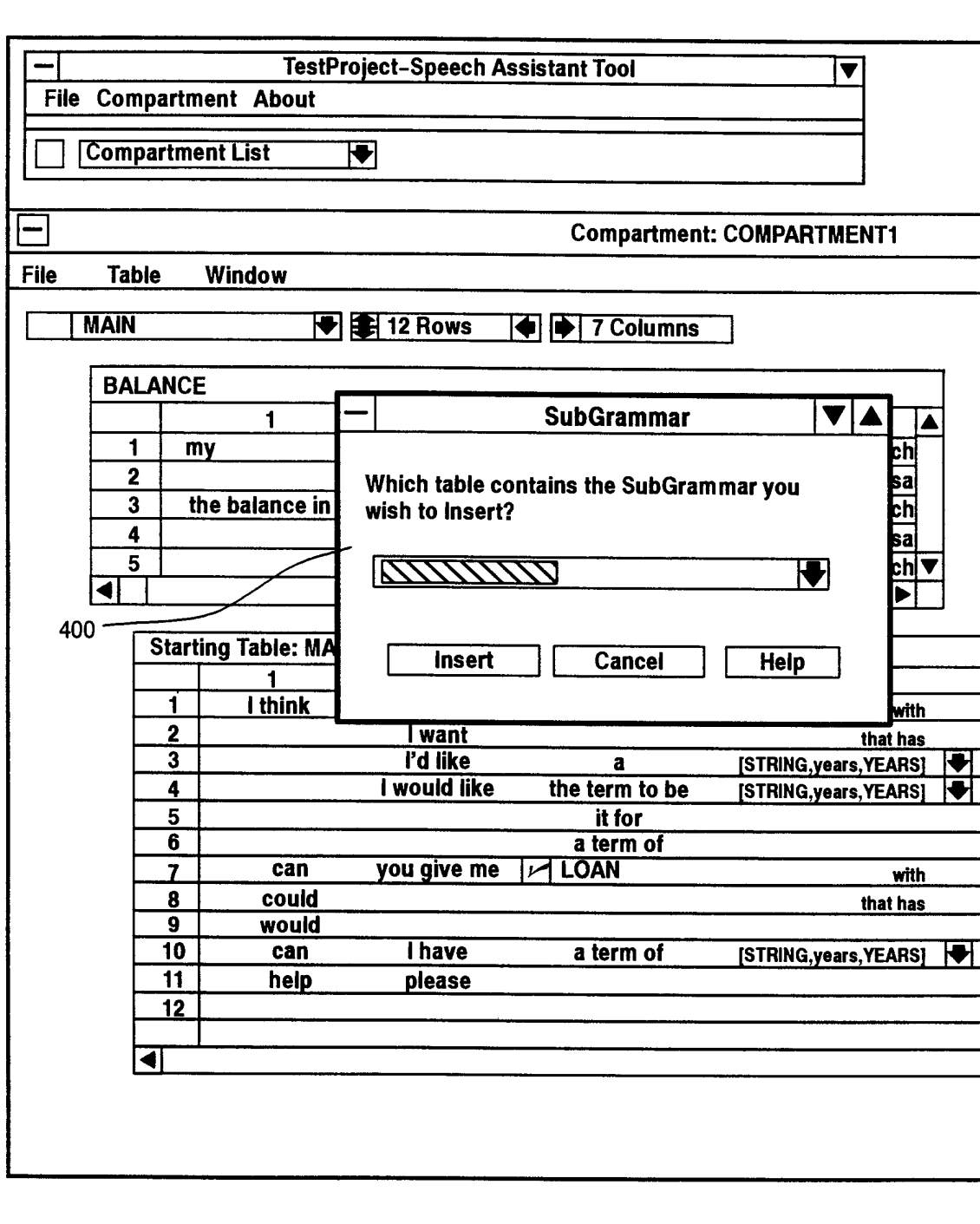
Figure 4a

Fig. 4a Fig. 4a

8 / 21

<div style="display: inline-block; border: 1px solid black; padding: 2px;">▼</div> <div style="display: inline-block; border: 1px solid black; padding: 2px;">▲</div>																																						
Hints																																						
<div style="display: flex; justify-content: space-around;"><div style="border: 1px solid black; padding: 5px; width: 45%;"><table border="1" style="width: 100%; border-collapse: collapse;"><tr><th colspan="3">LOAN</th></tr><tr><td style="width: 5%;"></td><td style="width: 85%;"></td><td style="width: 10%; text-align: center;">▲</td></tr><tr><td style="text-align: center;">1</td><td>a mortgage</td><td></td></tr><tr><td style="text-align: center;">2</td><td>a loan</td><td></td></tr><tr><td style="text-align: center;">3</td><td></td><td style="text-align: center;">▼</td></tr><tr><td style="text-align: center;">◀</td><td></td><td style="text-align: center;">▶</td></tr></table></div><div style="border: 1px solid black; padding: 5px; width: 45%;"><table border="1" style="width: 100%; border-collapse: collapse;"><tr><th colspan="3">YEARS</th></tr><tr><td style="width: 5%;"></td><td style="width: 85%;"></td><td style="width: 10%; text-align: center;">▲</td></tr><tr><td style="text-align: center;">1</td><td>20</td><td></td></tr><tr><td style="text-align: center;">2</td><td>30</td><td></td></tr><tr><td style="text-align: center;">3</td><td>60</td><td style="text-align: center;">▼</td></tr><tr><td style="text-align: center;">◀</td><td></td><td style="text-align: center;">▶</td></tr></table></div></div>			LOAN					▲	1	a mortgage		2	a loan		3		▼	◀		▶	YEARS					▲	1	20		2	30		3	60	▼	◀		▶
LOAN																																						
		▲																																				
1	a mortgage																																					
2	a loan																																					
3		▼																																				
◀		▶																																				
YEARS																																						
		▲																																				
1	20																																					
2	30																																					
3	60	▼																																				
◀		▶																																				
<table border="1" style="width: 100%; border-collapse: collapse;"><tr><td colspan="3"></td></tr><tr><td style="width: 33%; text-align: center;">5</td><td style="width: 33%; text-align: center;">6</td><td style="width: 34%; text-align: center;">TOKEN ▲</td></tr><tr><td>[STRING,years,YEARS] ▼</td><td>year term</td><td>Myyears</td></tr><tr><td colspan="3">year term</td></tr><tr><td><...></td><td>▼</td><td>years</td></tr><tr><td colspan="3">[STRING,years,YEARS] ▼</td></tr><tr><td colspan="3">years Tyears</td></tr><tr><td colspan="3">help</td></tr><tr><td colspan="3" style="text-align: right;">▼</td></tr><tr><td colspan="3" style="text-align: right;">▶</td></tr></table>						5	6	TOKEN ▲	[STRING,years,YEARS] ▼	year term	Myyears	year term			<...>	▼	years	[STRING,years,YEARS] ▼			years Tyears			help			▼			▶								
5	6	TOKEN ▲																																				
[STRING,years,YEARS] ▼	year term	Myyears																																				
year term																																						
<...>	▼	years																																				
[STRING,years,YEARS] ▼																																						
years Tyears																																						
help																																						
▼																																						
▶																																						

Figure 4a**Fig. 4a****Fig. 4a**

—		▼	▲
<p>Select An Option</p> <div> !Ignore !Class !Spelling !Export !Import !Optional </div>		<p>Lernout & Hauspie</p> <p>Selected ASR Option</p> <div>!Ignore</div> <p>Enter ASR Argument</p> <div></div> <div>Insert</div> <div>Cancel</div>	

Figure 4b

10/21

TestProject-Speech Assistant Tool

File Compartment About

☐

Compartment List

Compartment: COMPARTMENT1

File Table Window

MAIN

12 Rows

7 Columns

BALANCE

	1	2	3	4
1	my	checking	balance	
2		savings	balance	
3	the balance in my	checking	account	
4		savings	account	
5		checking	account	

Starting Table: MAIN

	1	2	3
1	I think	I need	<input checked="" type="checkbox"/> LOAN
2		I want	
3		I'd like	a
4		I would like	the term to be
5			it for
6			a term of
7	can	you give me	<input checked="" type="checkbox"/> LOAN
8	could		
9	would		
10	can	I have	a term of
11	help	please	
12			

Figure 4c

Fig. 4c

Fig. 4c

11/21

The image shows a software interface with a 'Hints' dialog box. The dialog box has a title bar with a minus sign and a 'Variable' tab. Inside the dialog, there is a section titled 'Select a Date Type' with a list of options: Date (American Format), Date (European Format), Currency, Currency (discrete), Time, Time (Military), Cardinal Integer (6 digit), Discrete Integer (10 digit), String (User defined), Cardinal Integer (User defined), and Discrete Integer (User defined). Below this list are 'Insert', 'Cancel', and 'Help' buttons. To the right of the list, there are four input fields: 'Variable Name:' with 'years' entered, 'Table Reference:' with 'Years' entered, 'If the user utters:' with '<value> from Table Reference' entered, and '...the SAI returns:' with '<value>' entered. Below these fields is a label '...with a maximum number of:' followed by an input field containing '60 characters'. A pointer labeled '420' points to the 'Table Reference' field. The dialog box is overlaid on a larger window. The background window has a table with columns and rows. The table has a header row with 'with' and '[STRING,years,YEARS]' and a body row with 'that has' and '[STRING,years,YEARS]'. There are also buttons for 'Tyears' and 'help'.

Figure 4c

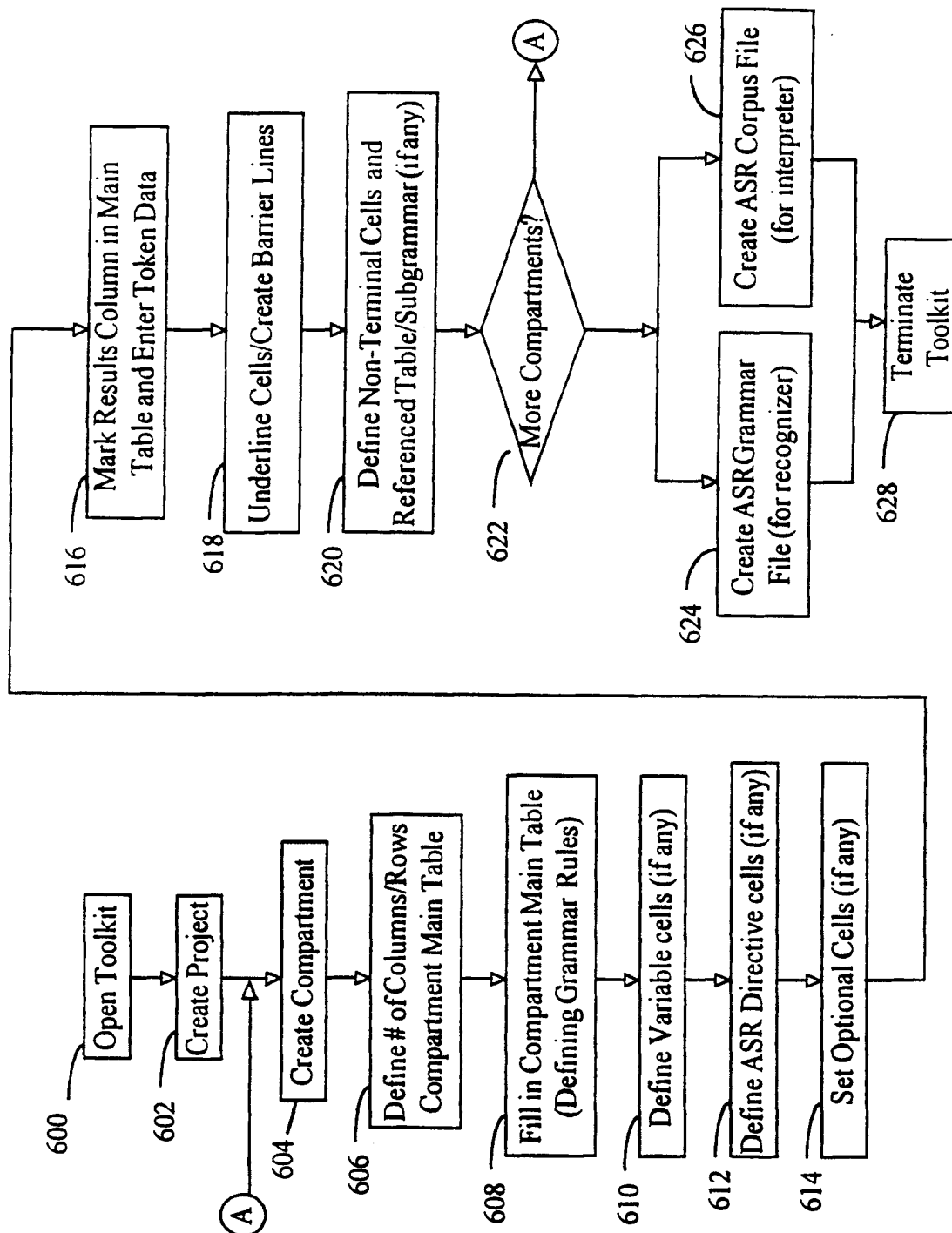
Fig. 4c

Fig. 4c

12/21

Name	Accepts	Returns
TIME1	one o'clock p.m.; thirteen hundred hours	1:00 PM (normal format)
TIME2	one o'clock p.m.; thirteen hundred hours	13:00 (European/military format)
INTEGER1	one hundred twenty three	digits
INTEGER2	one two three . . . (ten digits)	digits
CURRENCY1	twenty three dollars and fifteen cents	23.15
CURRENCY2	two three dollars and fifteen cents	23.15
DATE1	march first nineteen ninety seven; the first of march	03/01/1997 (American format)
DATE2	march first nineteen ninety seven; the first of march	01/03/1997 (European format)

Figure 5

**Figure 6**

SAT - Annotated Corpus			
ProjectTestProject		8/14/97 10:53:02 AM	
Compartment Name	Compartment Corpus File Name	Creation Date	
<input checked="" type="checkbox"/> COMPARTMEN...	C:\SAT 1.2.x\Projects\TestProject\Cor...	8/14/97 10:53:00 AM	
<div style="border: 1px solid black; padding: 5px; margin-top: 10px;"> Processing Results The annotated Corpus has been saved in the following files: C:\SAT 1.2.x\Projects\TestProject\SA\TestProjectAnswers.txt C:\SAT 1.2.x\Projects\TestProject\SA\TestProjectCorpus.txt </div>			
<div style="border: 1px solid black; display: inline-block; padding: 5px 15px; margin: 5px;">Create</div> <div style="border: 1px solid black; display: inline-block; padding: 5px 15px; margin: 5px 20px;">Close</div> <div style="border: 1px solid black; display: inline-block; padding: 5px 15px; margin: 5px;">Help</div>			

Figure 7

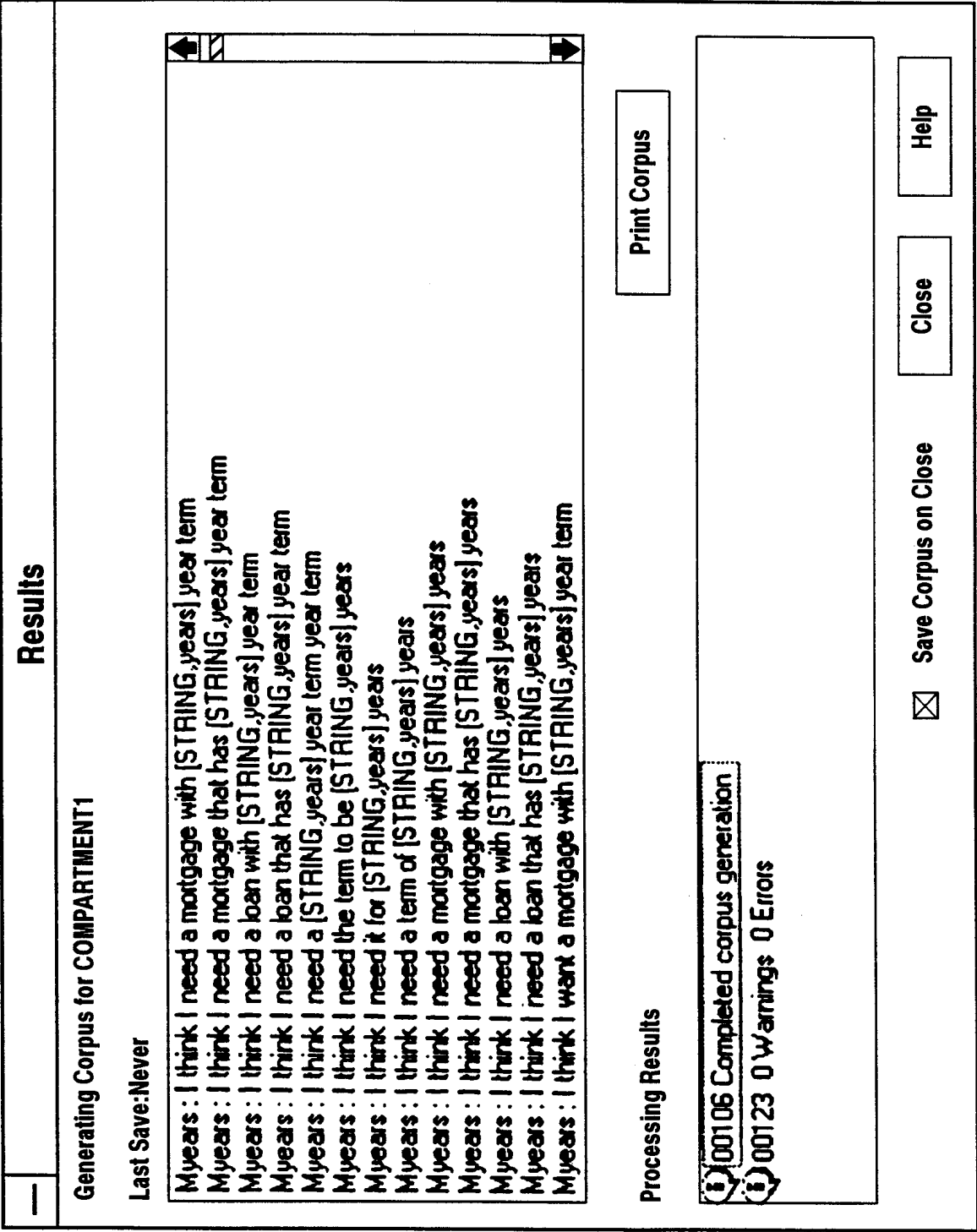


Figure 8a

Results

Generating Grammar for COMPARTMENT1

Last Saved: 8/14/97 10:37:52AM

!grammar MAIN:
/* language Eng_NZ : */
!start <MAIN> ;

<MAIN> : { { { ["I think"? | ("can"? | "could" | "would") | "I need" | "I want" | "I'd like" | "I
[<LOAN> { "with" | "that has" } <YEARS> } | { "a" <YEARS> "year term" } } "year term"
"a term of" } <YEARS> <...> } | { <LOAN> { "with" | "that has" } <YEARS> } } "years" } }
<YEARS> "years" } | { "help" ("please"?) } } ;

<LOAN> : { "a mortgage" | "a loan" } ;

<YEARS> : { "20" | "30" | "60" } ;

Processing Results

00123 Generating grammar MAIN

00123 Generating grammar LOAN

00123 Generating grammar YEARS

00123 Generating grammar BALANCE

00108 Successfully completed grammar generation

00123 0 Warnings 0 Errors

Print Corpus

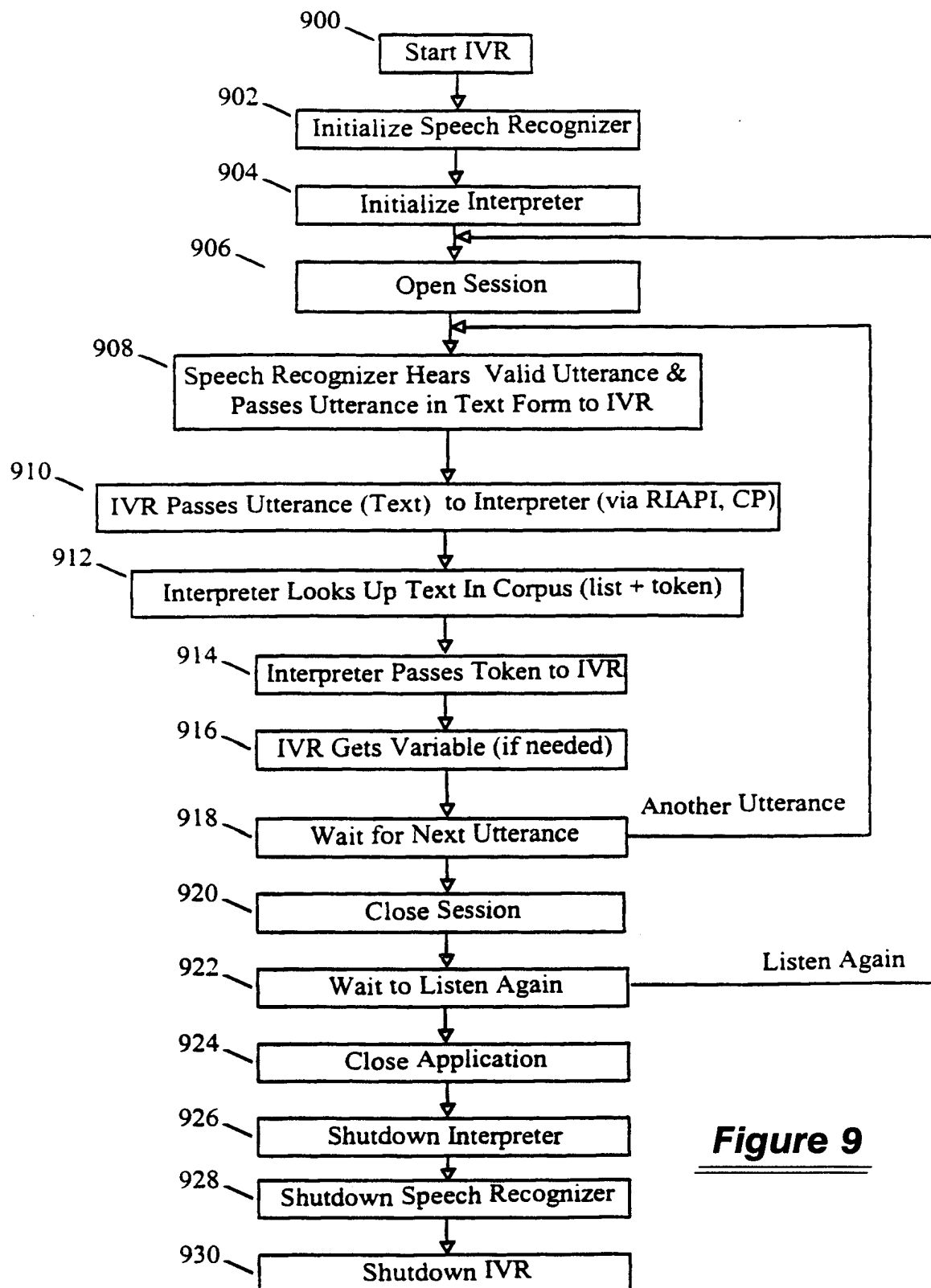
☒ Save Grammar on Close

Close

Help

Figure 8b

17/21

**Figure 9**

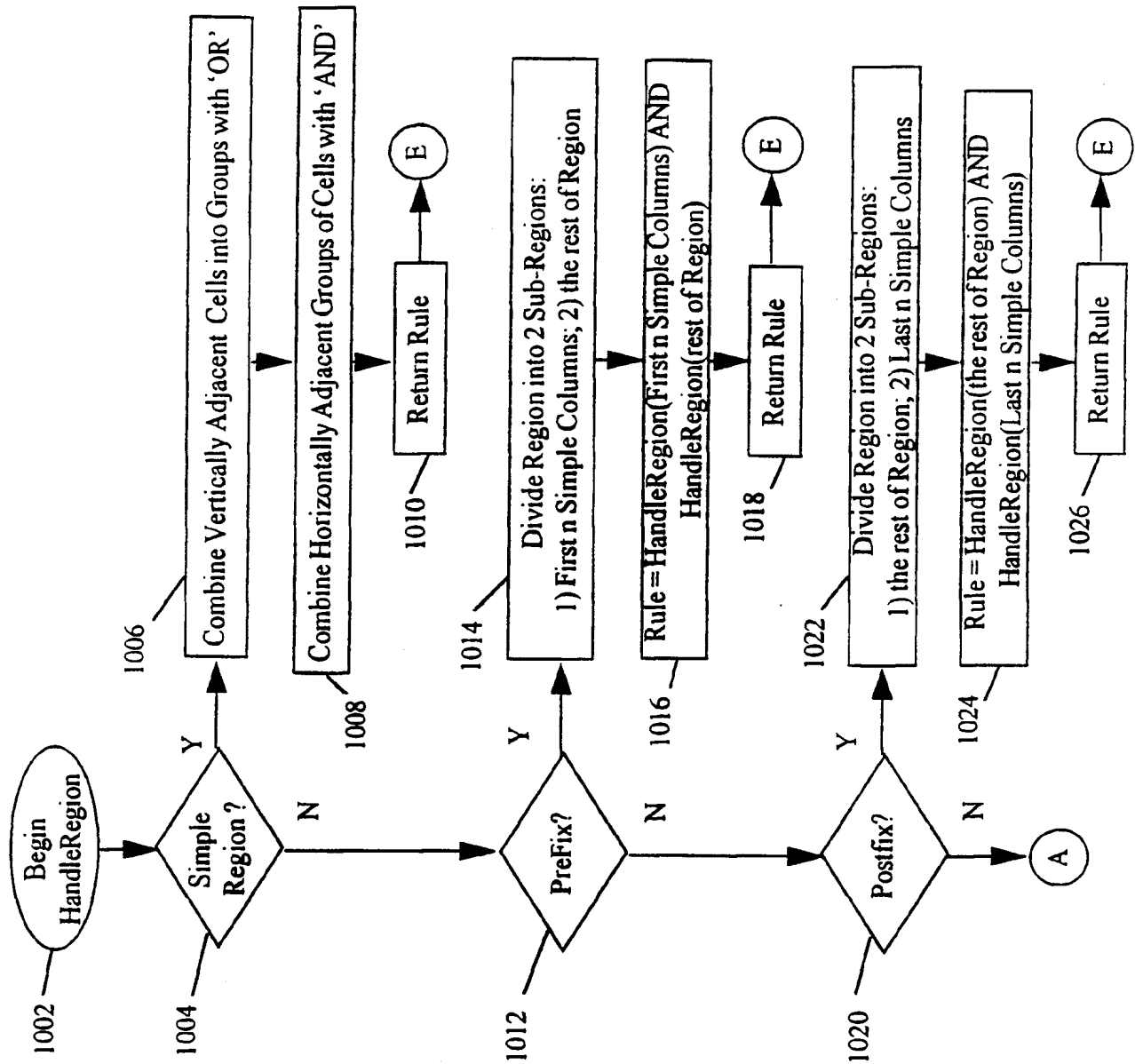
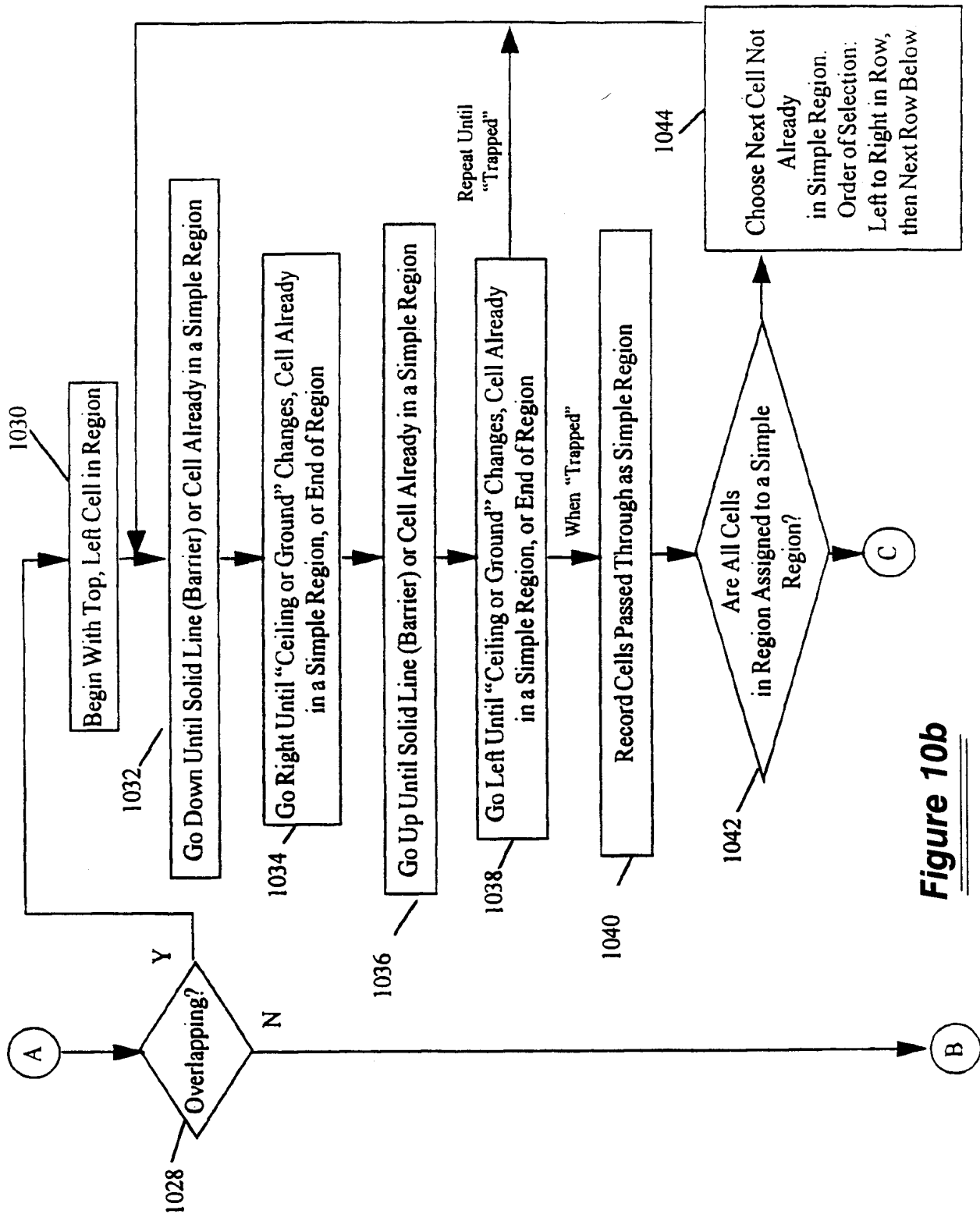
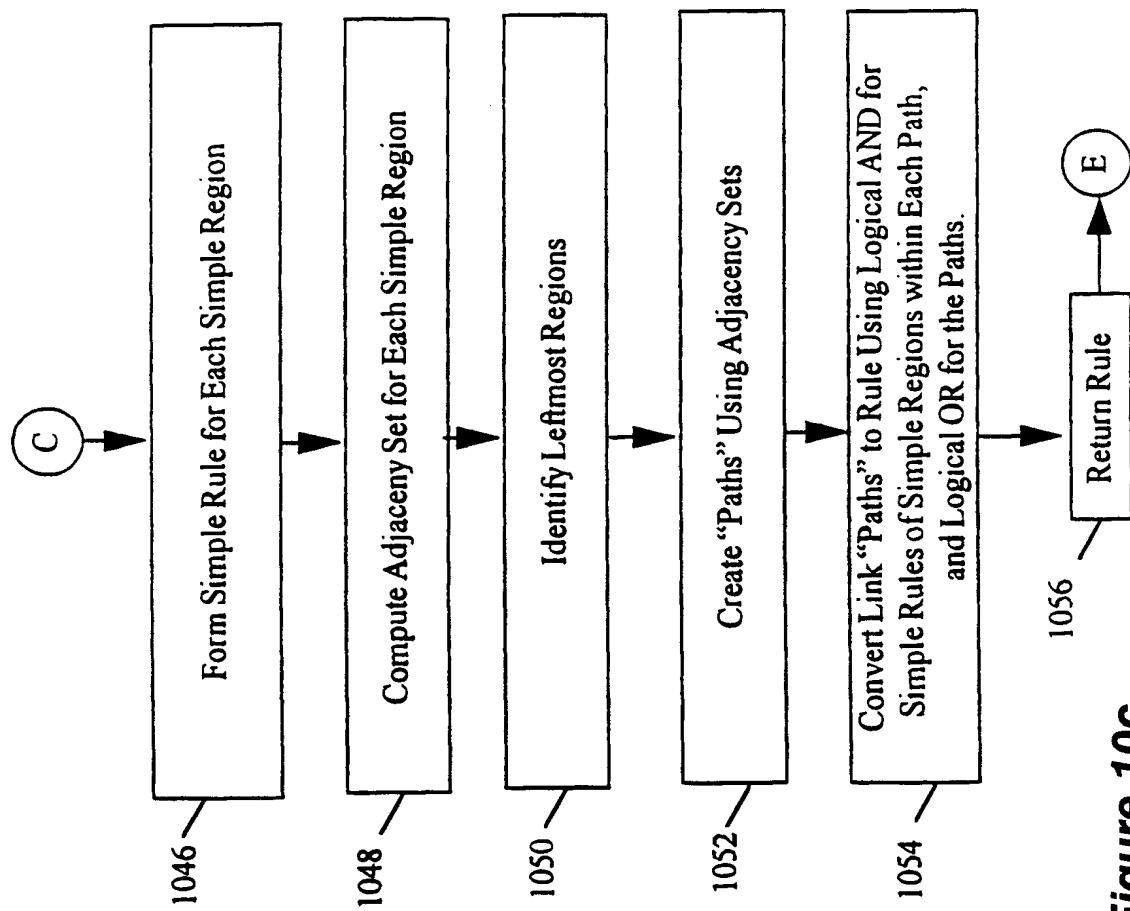
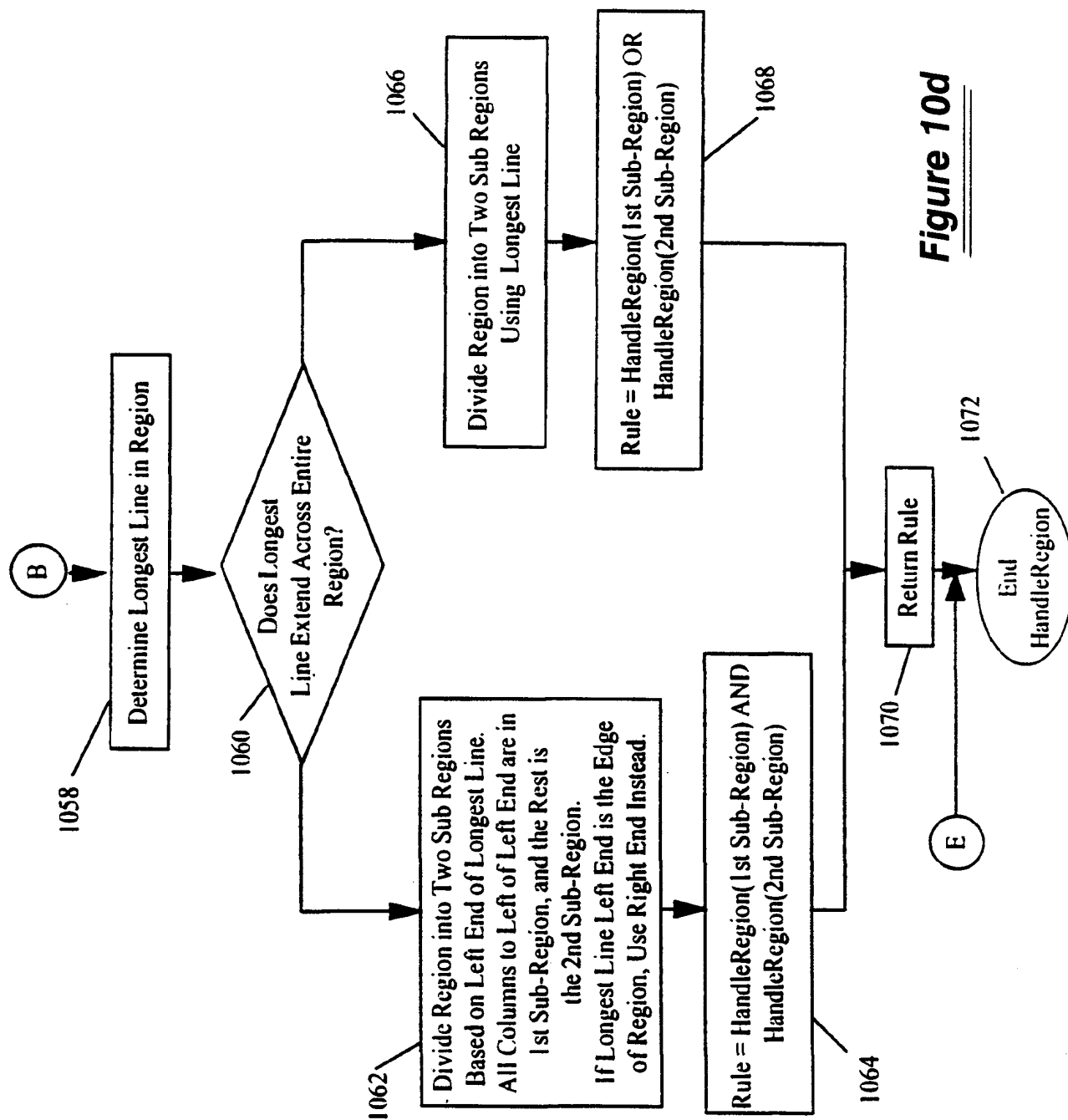


Figure 10a

**Figure 10b**

20/21

**Figure 10c**

**Figure 10d**

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 98/19432

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 G06F17/27

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 6 G06F G10L

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	W0 96 26484 A (EGILSSON AGUST S) 29 August 1996 see page 2, line 24 - page 5, line 27 see figure 8 ---	1-17
A	EP 0 681 284 A (SUN MICROSYSTEMS INC) 8 November 1995 see page 4, line 33 - page 5, line 3 see figure 1 ---	1-17
A	US 5 544 298 A (KANAVY WALTER J ET AL) 6 August 1996 see column 2, line 48 - line 65 -----	1-17

☐ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

° Special categories of cited documents :

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier document but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

- "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- "&" document member of the same patent family

Date of the actual completion of the international search

29 January 1999

Date of mailing of the international search report

05/02/1999

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016

Authorized officer

Pedersen, N

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/US 98/19432

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9626484 A	29-08-1996	AU 4783796 A CA 2214972 A DE 69600794 D EP 0811193 A	11-09-1996 29-08-1996 19-11-1998 10-12-1997
EP 0681284 A	08-11-1995	US 5642519 A AU 686324 B AU 1657195 A CA 2147917 A JP 8063181 A	24-06-1997 05-02-1998 09-11-1995 30-10-1995 08-03-1996
US 5544298 A	06-08-1996	NONE	